

WHAT IS CLAIMED IS:

1. An image processing apparatus which compresses an input multilevel image, comprising:

binarizing means for binarizing the multilevel
5 image;

first region specifying means for specifying a position of a character region in the binary image binarized by said binarizing means, and determining whether or not the specified character region is an
10 inverted character region;

second region specifying means for specifying a position of a unit character region in the character region specified by said first region specifying means;

reducing means for reducing the multilevel image;
15 representative color extracting means for extracting a representative color of the unit character region on the basis of region type information indicating whether or not the character region specified by said first region specifying means is an inverted
20 character region, position information of the unit character region specified by said second region specifying means, and the reduced multilevel image;

padding means for padding a unit character region in the character region of the reduced multilevel image
25 with a neighboring color on the basis of the position information of the character region specified by said first region specifying means;

first compression means for compressing the padded reduced multilevel image created by said padding means;

second compression means for compressing a
5 partial binary image corresponding to the character region specified by said first region specifying means; and

output means for outputting compressed data containing the position information specified by said
10 first and second region specifying means, first and second compressed codes created by said first and second compression means, and representative color information of the unit character region extracted by said representative color extracting means.

15 2. The apparatus according to claim 1, wherein said representative color extracting means comprises

average color calculating means for calculating an average color of the unit character regions by using the position information of the unit character region
20 specified by said second region specifying means and referring to color information of the reduced multilevel image,

first histogram creating means for creating a first histogram from the average color of the unit
25 character regions created by said average color calculating means,

extracting means for extracting a candidate

representative color serving as a representative color candidate of the unit character region on the basis of the first histogram, and

color assigning means for assigning the candidate
5 representative color to a representative color of the unit character on the basis of region type information indicating whether or not the character region specified by said first region specifying means is an inverted character region, the candidate representative
10 color extracted by said extracting means, and the average color of the unit character regions.

3. The apparatus according to claim 2, wherein said binarizing means comprises

second histogram creating means for creating a
15 second histogram on the basis of the multilevel image, and

binarization threshold calculating means for calculating a binarization threshold for binarizing the multilevel image on the basis of the second histogram.

20 4. The apparatus according to claim 1, wherein said first region specifying means specifies regions, which are specified by contour line tracking of pixels having a predetermined value in the binary image, as a character region or a natural image region, specifies
25 whether the character region is an inverted character region, and specifies attribute information indicating position information, size information, and a type of

each of the regions.

5. The apparatus according to claim 1, wherein said second region specifying means specifies, as a unit character region, a set of pixels having a
5 predetermined value in the character region specified by said first region specifying means, and specifies position information and size information of the unit character region.

6. The apparatus according to claim 2, wherein said
10 average color calculating means calculates an average color of the unit character regions by using position information of a unit character region specified by said second region specifying means and referring to color information of the reduced multilevel image
15 corresponding to pixels having a predetermined value in the unit character region.

7. The apparatus according to claim 2, wherein said first histogram creating means creates a color space histogram in a character region constituted by the unit
20 character regions as a first histogram by using an average color of the unit character regions which is calculated by said average calculating means.

8. The apparatus according to claim 7, wherein said extracting means extracts candidate representative
25 colors as candidates of representative colors of the unit character regions in descending order of output frequency of color distributions in the color spatial

histogram.

9. The apparatus according to claim 8, wherein said color assigning means calculates a distance between a candidate representative color extracted by said
5 extracting means and an average color of the unit character regions on a color space, and assigns the candidate representative color as a representative color of the unit character regions if the distance falls within a predetermined range.

10. The apparatus according to claim 8, wherein said color assigning means calculates a distance between a candidate representative color extracted by said
10 extracting means and an average color of the unit character regions on a color space, and assigns a predetermined color as a representative color of the
15 unit character regions if the distance falls within a predetermined range and the character region constituted by the unit character regions is an inverted character region.

20 11. The apparatus according to claim 1, wherein said reducing means converts a resolution of the multilevel image by using a resolution conversion parameter.

12. The apparatus according to claim 1, wherein said padding means pads a unit character region in a
25 character region of the reduced multilevel image corresponding to a position corresponding to a reduction ratio set by said reducing means on the basis

of position information of the character region specified by said first region specifying means.

13. The apparatus according to claim 1, wherein said first compression means complies with JPEG compression.

5 14. The apparatus according to claim 1, wherein said second compression means complies with MMR compression.

15. The apparatus according to claim 1, further comprising third compression means for losslessly compressing the compressed data.

10 16. An image processing apparatus which decompresses compressed data, comprising:

input means for inputting compressed data containing a first compressed code obtained by compressing a character region or inverted character
15 region in a binary image obtained by binarizing a multilevel image, a second compressed code obtained by compressing a reduced multilevel image obtained by padding a unit character region in a character region or inverted character region in the reduced multilevel
20 image of the multilevel image with a neighboring color, position information of the character region or inverted character region, and representative color information of a unit character region in the character region or inverted character region;

25 first decompressing means for decompressing the first compressed code in the compressed data;

second decompressing means for decompressing the

second compressed code in the compressed data;

enlarging means for enlarging a reduced multilevel image decompressed by said second decompressing means; and

5 combining means for combining a binary image decompressed by said decompressing means and a multilevel image enlarged by said enlarging means.

17. An image processing method of compressing an input multilevel image, comprising:

10 a binarizing step of binarizing the multilevel image;

a first region specifying step of specifying a position of a character region in the binary image binarized in the binarizing step, and determining
15 whether or not the specified character region is an inverted character region;

a second region specifying step of specifying a position of a unit character region in the character region specified in the first region specifying step;

20 a reducing step of reducing the multilevel image;

a representative color extracting step of extracting a representative color of the unit character region on the basis of region type information indicating whether or not the character region specified
25 in the first region specifying step is an inverted character region, position information of the unit character region specified in the second region

specifying step, and the reduced multilevel image;

a padding step of padding a unit character region
in the character region of the reduced multilevel image
with a neighboring color on the basis of the position
5 information of the character region specified in the
first region specifying step;

a first compression step of compressing the
padded reduced multilevel image created in the padding
step;

10 a second compression step of compressing a
partial binary image corresponding to the character
region specified in the first region specifying step;
and

an output step of outputting compressed data
15 containing the position information specified in the
first and second region specifying steps, first and
second compressed codes created in the first and second
compression steps, and representative color information
of the unit character region extracted in the
20 representative color extracting step.

18. An image processing method of decompressing
compressed data, comprising:

an input step of inputting compressed data
containing a first compressed code obtained by
25 compressing a character region or inverted character
region in a binary image obtained by binarizing a
multilevel image, a second compressed code obtained by

compressing a reduced multilevel image obtained by padding a unit character region in a character region or inverted character region in the reduced multilevel image of the multilevel image with a neighboring color,
5 position information of the character region or inverted character region, and representative color information of a unit character region in the character region or inverted character region;

a first decompressing step of decompressing the
10 first compressed code in the compressed data;

a second decompressing step of decompressing the second compressed code in the compressed data;

an enlarging step of enlarging a reduced multilevel image decompressed in the second
15 decompressing step; and

a combining step of combining a binary image decompressed in the decompressing step and a multilevel image enlarged in the enlarging step.

19. A program which realizes image compression of
20 compressing an input multilevel image, comprising:

a program code for a binarizing step of binarizing the multilevel image;

a program code for a first region specifying step of specifying a position of a character region in the
25 binary image binarized in the binarizing step, and determining whether or not the specified character region is an inverted character region;

a program code for a second region specifying step of specifying a position of a unit character region in the character region specified in the first region specifying step;

5 a program code for a reducing step of reducing the multilevel image;

a program code for a representative color extracting step of extracting a representative color of the unit character region on the basis of region type
10 information indicting whether or not the character region specified in the first region specifying step is an inverted character region, position information of the unit character region specified in the second region specifying step, and the reduced multilevel
15 image;

a program code for a padding step of padding a unit character region in the character region of the reduced multilevel image with a neighboring color on the basis of the position information of the character
20 region specified in the first region specifying step;

a program code for a first compression step of compressing the padded reduced multilevel image created in the padding step;

a program code for a second compression step of
25 compressing a partial binary image corresponding to the character region specified in the first region specifying step; and

a program code for an output step of outputting compressed data containing the position information specified in the first and second region specifying steps, first and second compressed codes created in the first and second compression steps, and representative color information of the unit character region extracted in the representative color extracting step.

20. A program which realizes image decompression of decompressing compressed data, comprising:

10 a program for an input step of inputting compressed data containing a first compressed code obtained by compressing a character region or inverted character region in a binary image obtained by binarizing a multilevel image, a second compressed code
15 obtained by compressing a reduced multilevel image obtained by padding a unit character region in a character region or inverted character region in the reduced multilevel image of the multilevel image with a neighboring color, position information of the
20 character region or inverted character region, and representative color information of a unit character region in the character region or inverted character region;

a program for a first decompressing step of
25 decompressing the first compressed code in the compressed data;

a program for a second decompressing step of

decompressing the second compressed code in the compressed data;

a program for an enlarging step of enlarging a reduced multilevel image decompressed in the second decompressing step; and

a program for a combining step of combining a binary image decompressed in the decompressing step and a multilevel image enlarged in the enlarging step.

21. An image processing apparatus which compresses an input multilevel image, comprising:

binarizing means for binarizing the multilevel image;

first region specifying means for specifying a position of a character region in the binary image binarized by said binarizing means, and determining whether or not the specified character region is an inverted character region;

second region specifying means for specifying a position of a unit character region in the character region specified by said first region specifying means;

representative color extracting means for extracting a representative color of the unit character region on the basis of region type information indicating whether or not the character region specified by said first region specifying means is an inverted character region; position information of the unit character region specified by said second region

specifying means, and the multilevel image.

22. The apparatus according to claim 21, wherein said extracting means extracts a representative color of the inverted character region, after the inverted character region is applied by an inversion processing.

23. An image processing method of compressing an input multilevel image, comprising:

a binarizing step of binarizing the multilevel image;

10 a first region specifying step of specifying a position of a character region in the binary image binarized in the binarizing step, and determining whether or not the specified character region is an inverted character region;

15 a second region specifying step of specifying a position of a unit character region in the character region specified in the first region specifying step;

a representative color extracting step of extracting a representative color of the unit character region on the basis of region type information indicting whether or not the character region specified in the first region specifying step is an inverted character region, position information of the unit character region specified in the second region specifying step, and the multilevel image.

24. The method according to claim 23, wherein said representative color extracting step extracts a

representative color of the inverted character region,
after the inverted character region is applied by an
inversion processing.